Informatics in Medicine Unlocked 22 (2021) 100503

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: http://www.elsevier.com/locate/imu

Transcription-translation error: *In-silico* investigation of the structural and functional impact of deleterious single nucleotide polymorphisms in *GULP1* gene

Opeyemi S. Soremekun [a,e], Chisom Ezenwa [b,e], Mahmoud Soliman [a], Tinashe Chikowore [c,d], Oyekanmi Nashiru [b,**], Segun Fatumo [b,e,f,*]

[a] Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban, 4001, South Africa
[b] Centre for Genomics Research and Innovation, National Biotechnology Agency, Abuja, Nigeria
[c] MRC/Wits Developmental Pathways for Health Research Unit, Department of Pediatrics, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[d] Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
[e] The African Computational Genomics (TACG) Research Group, MRC/UVRI LSHTM Uganda Research Unit, Entebbe, Uganda
[f] Department of Non-communicable Disease Epidemiology (NCDE), London School of Hygiene and Tropical Medicine, London, United Kingdom

ARTICLE INFO

Keywords:
Nonsynonymous single nucleotide polymorphisms (nsSNPs)
GULP1
SNPinformatics
MD Simulation

ABSTRACT

Nonsynonymous single nucleotide polymorphisms (nsSNPs) are one of the most common forms of mutations known to disrupt the product of translation thereby altering the protein structure-function relationship. *GULP1* (PTB domain-containing engulfment adaptor protein 1) is an evolutionarily conserved adaptor protein that has been associated with glycated hemoglobin (HbA1c) in Genome-Wide Association Studies (GWAS). In order to understand the role of *GULP1* in the etiology of diabetes, it is important to study some functional nsSNPs present within the *GULP1* protein. We, therefore, used a SNPinformatics approach to retrieve, classify, and determine the stability effect of some nsSNPs. Y27C, G142D, A144T, and Y149C were jointly predicted by the pathogenic-classifying tools to be disease-causing, however, only G142D, A144T, and Y149C had their structural architecture perturbed as predicted by I-MUTANT and MuPro. Interestingly, G142D and Y149C occur at positions 142 and 149 of *GULP1* which coincidentally are found within the binding site of *GULP1*. Protein-Protein interaction analysis also revealed that *GULP1* interacted with 10 proteins such as Cell division cycle 5-like protein (CDC5L), ADP-ribosylation factor 6 (ARF6), Arf-GAP with coiled-coil (ACAP1), and Multiple epidermal growth factor-like domains protein 10 (MEGF10), etc. Taken together, rs1357922096, rs1264999716, and rs128246649 could be used as genetic biomarkers for the diagnosis of diabetes. However, being a computational study, these nsSNPs require experimental validation to explore their metabolic involvement in the pathogenesis of diseases.

1. Introduction

The advancements in genetic technology, have led to the identification of several genetic variations within the human genome, this is however not without some challenges. Single Nucleotide Polymorphism (SNP) is the most abundant form of variation in the human genome and range from 3 to 5 million in number [1]. SNPs have been discovered to contribute to the onset of diseases, hence, they are used as genetic markers for understanding the etiology of diseases, however, some SNPs are neutral, that is, they are not disease-causing [2]. SNPs that alter the primary sequence of an amino acid are referred to as nonsynonymous single nucleotide polymorphisms (nsSNPs). Due to the changes they elicit on the amino acid chain, the product of translation is affected, therefore distorting the functions of proteins [3], and possibly drug metabolism and absorption [4]. These SNPs have been extensively studied through Genome-Wide Association Study (GWAS) or family-based study's [4].

GULP1 (PTB domain-containing engulfment adaptor protein 1) is an evolutionarily conserved adaptor protein necessary for phagocytosing apoptotic cells through phagocytosis [5]. It contains several

* Corresponding author. The African Computational Genomics (TACG) Research Group, MRC/UVRI and LSHTM, Entebbe, Uganda.
** Corresponding author.
E-mail addresses: oyekan.nash@gmail.com (O. Nashiru), segun.fatumo@lshtm.ac.uk (S. Fatumo).

https://doi.org/10.1016/j.imu.2020.100503
Received 21 October 2020; Received in revised form 14 December 2020; Accepted 14 December 2020
Available online 24 December 2020

protein-interacting domains and regions such as the N-terminal phosphotyrosine-binding (PTB) domain, leucine zipper domain, and a proline/serine rich-domain [6,7]. They regulate cellular cholesterol and glycosphingolipid translocation [5]. Aberration in the structure of *GULP1* has been reported in several diseases such as schizophrenia [8], arthritis [9], and cancer [10].

In our previous GWAS analysis, we identified the association between *GULP1* and glycated hemoglobin (HbA1c) [11]. HbA1c is a marker used for diagnosing diabetes and is often used in the estimation of the average glucose level for last three months in individuals to keep track of diabetic vascular damage [12]. HbA1c is not only used for the diagnosis of diabetes mellitus, it can be used in the diagnosis of diabetes-induced pathologies such as retinopathy, nephropathy and neuropathy. However, HbA1C is also not fully utilised for diagnosis in Africa due to concerns around it being confounded by infections and iron deficiency. Taking into consideration the role GULP1 plays in the diagnosis of diseases, it is important to study the functional consequence of variants present within this gene. We therefore carried out a SNPinformatics investigation on SNP data retrieved from NCBI-SNPs database employing pathogenic predicting tools such as PolyPhen, PANTHER, SNP&GO, PhD-SNP, etc. We further investigated the conservation and stability of the predicted pathogenic nsSNPs. Furthermore, structural examination and visualization of the pathogenic SNPs were carried out using molecular dynamic simulation and CHIMERA respectively. The investigational approach (Fig. 1) employed in this study has an advantage of enabling the quick and cheap screening of functionally important variants which will need to be validated by experimental procedures.

2. Materials and methods

2.1. GULP1 structural elucidation

To characterize the 3-dimensional structure of *GULP1*, we used a computational structure based technique facilitated by Iterative Threading Assembly Refinement (I-TASSER) algorithm [13]. I-TASSER algorithm is trained with thousands of protein models that serve as templates for building structures of native proteins or proteins that lack 3D structure [13]. COACH [14] was employed to determine the potential binding pockets of the *GULP1* protein. COACH combines the binding site predictions of TM-SITE [15], FINDSITE [16], and ConCavity [17] to predict potential binding pockets. Amino acid changes were exerted in *GULP1* 3D structure by using "swapaa" command line in CHIMERA [18].

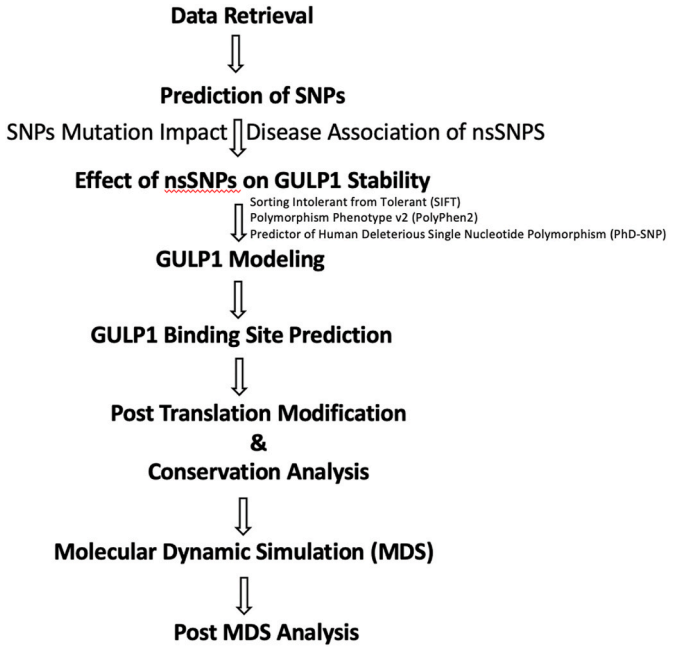

Fig. 1. Schematic workflow of the steps undertaken in this study.

2.2. GULP1 nsSNP dataset retrieval and prediction of functional impact of nsSNPs

UniProt database [19] was used as the source for retrieving the FASTA sequence of GULP1 gene (Q9UBP9) while the *GULP1* nsSNPs dataset was downloaded from the dbSNPs database [20] and further cross-validated with the gnomAD browser [21] and Ensembl browser [22]. To access the pathogenicity of *GULP1* nsSNPs, SIFT [23], Polyphen2 [24], and PhD-SNP [25] were engaged. SIFT accesses the pathogenic status of a SNP by using a ranking score known as tolerance index (TI) score [23]. Mutations with a TI score <0.05 are regarded as being harmful whereas those with TI values $>$ 0.05 are referred to as non-pathogenic [23]. PolyPhen moreover investigates the off chance probability that the change is found in the evolutionary conserved domain, using the PSIC score [24]. PhD-SNP employs a trained dataset to examine the pathogenic nature of an amino acid substitution as either disease-causing or non-disease causing [25]. SNPs&GO is a very precise and accurate server which projects disease-related amino acid substitution with 82% accuracy [26]. PMUT is a tool that annotates and predicts whether an amino acid substitution at a position is pathological or non-pathological. PMUT uses different types of backend algorithms to describe a mutation and neural networks to process the data [27]. SNAP2 is a tool trained based on a neural network, it differentiates between effect and non-synonymous variation by using information from the sequence and variants. It uses a prediction score that ranges from -100 to $+100$ signifying a strong non-synonymous prediction and strong synonymous prediction respectively [28].

2.3. Effect of mutation on GULP1 structural stability

The probability of an nsSNP altering the strength of a protein is very high, it either increases or decreases the stability of the protein. To evaluate the effect of nsSNP on *GULP1* stability, three tools (I-Mutant [29] and MUpro [30]) were used. I-Mutant is a support vector machine tools that automatically predicts protein stability differences upon amino acid substitution. I-Mutant is optimized to predict protein stability using a protein structure or protein sequence. If the protein structure is used as a query, the accuracy of I-Mutant is 80% while it is 77% if protein sequence is used. The energy change (DDG) value which is the Gibbs free energy of the mutant minus Gibbs free energy of the wild protein in Kcal/mol. MUpro is a tool that uses the combination of two machine learning approaches support vector machines and neural networks to predict how single nucleotide mutation affects protein stability, it has an accuracy of 84%. A DDG score of less than 0 signifies that the amino acid substitution decreases protein stability while a score greater than 0 connotes increase protein stability.

2.4. GULP1 sequence conservation analysis and post-transcriptional modification sites (PTMs) prediction

Estimation of the conserved region present within *GULP1* was predicted with the aid of the ConSurf server [31]. ConSurf estimates the conservation value based on the evolutional relatedness between the proteins and its homologs. Post-transcription modification has been implicated in events involved in the pathogenesis of diseases such as signaling pathways, protein-protein interaction. Thus, the prediction of PTM helps to provide insight on the impact of variations in pathogenesis of diseases. We used Modpred and Musite to determine the PTM sites present in *GULP1*. Modpred predicts PTM sites using protein sequence.

2.5. Time-wise structural perturbatory effect of mutation on GULP1

We employed molecular dynamic simulation (MDS) to explore the structural perturbatory impact of mutation on *GULP1* using a simulation protocol previously reported [32,33]. GULPI proteins (wild and mutants) parametrization was carried out with the aid of FF14SB present in AMBER18 software. GULP1 parameter and topology coordinates were derived using the LEAP variant of AMBER18. A partial restrain, full restrain, heating, and equilibration [34] were then carried out. A production run of 100ns was then subsequently done [35]. The production run trajectories were examined with the aid of the CPPTRAJ variant of AMBER18 [36]. Visualization and amino acid substitution were facilitated by CHIMERA [18].

3. Results

3.1. GULP1 structure determination

The 3-dimensional structure of *GULP1* has been previously crystalized (6ITU) [37] and deposited in the Protein Data Bank (PDB). However, this crystalized structure is only a part of the whole *GULP1* protein. Hence, it was necessary to build a *GULP1* structure containing all the important domains. The FASTA sequence of *GULP1* was retrieved from the UniProt server [19] and consequently used as an input in I-TASSER, 6ITU and 3SUZ were used as templates. The modeled GULP1 protein has a confidence score (C-score), estimated TM-Score, and an estimated Root Mean Square Deviation (RMSD) [38] of −2.59, 0.41 ± 0.14, and 12.3 ± 4.3 Å respectively (Fig. 2A). GULP1 was validated with the aid of Verify-3D, PROCHECK, and ERRAT. Investigation using these tools revealed that the modeled GULP1 has high structural integrity and could be used for further downstream bioinformatic analysis. Binding site characterization using COACH identified 22 residues (residues P40, K41, T43, L95, H96, R97, I98, S99, F100, C101, A102, D103, K105, K116, H123, E135, T138, L139, G142, F145, Y149, and F152) at a C-score of 0.43, COACH predicted two out of the three predicted nsSNPs occur at positions (142 and 149) located within the active site.

3.2. Prediction of pathogenic SNPs and evaluating GULP1 protein change in stability

GULP1 SNPs retrieved from dbSNPs contains 628 SNPs, out of these, 207 were non-synonymous. The 207 nsSNPs were inputted into the PhD-SNP, PANTHER, PolyPhen, and SNPs&GO servers. Four nsSNPs (rs1446644508, rs1357922096, rs1264999716, and rs128246649) were predicted to be disease-causing by these four SNPs-pathogenic predicting tools (Table 1). Further analysis of these four SNPs with PMUT and SNAP2 revealed that they were also pathogenic and disease-causing. The four predicted SNPs were further analysed to evaluate their effect on *GULP1* stability. I-MUTANT and MuPro were employed to determine the change in protein instability. rs1446644508 was predicted by I-MUTANT and MuPro to increase and decrease GULP1 stability respectively. However, rs1357922096, rs1264999716, and rs128246649 were jointly predicted by I-MUTANT and MuPro to decrease stability of GULP1. Hence, theses SNPs were used for further analysis.

3.3. Evaluation of post translational modification sites and conservation analysis

Post Translational Modifications linked to our predicted nsSNPs were determined with the aid of the ModPred server by using the primary sequence of *GULP1* as input. ModPred predicted that rs128246649 is located in the proteolytic cleavage site, while rs1357922096, rs1264999716, and rs1446644508 had no PTM site. Conservation analysis is important in unraveling whether the nsSNPs are found in a conserved region or not. Analysis from the ConSurf server revealed that rs1446644508 and rs1357922096 are located in a highly conserved region of *GULP1* with both having a conversation score of 8. rs128246649 occur in an averagely conserved region with a conservation score of 5 (Table 2).

3.4. Time-wise structural perturbatory effect of mutation on GULP1

To explore the time-wise effect of the mutation on the mutant proteins relative to the wild type, we examined the Root Mean Square Deviation (RMSD), Radius of Gyration (RoG), Principal Component Analysis (PCA), and the number of hydrogen bonds in the proteins. As seen above in the stability estimation using I-MUTANT and MuPro, the mutation altered the stability of the protein by causing a decrease in the stability of the mutant. Likewise, the Cα backbone RMSD plot corroborates these findings, the wild protein displayed high stability throughout the simulation period, however, upon mutation, the instability of the proteins was markedly increased (Fig. 3A). Similarly, when the proteins were projected on two motional components, principal component 1 (PC1) and principal component 2 (PC2), the wild protein had little dispersion along the two principal components (PC1 and PC2), unlike the mutant proteins (G142D, A144T, and Y149C) that exhibited a highly dispersed motion along the two principal component, this is as a result of the instability conferred on the proteins by mutation (Fig. 3B) (see Fig. 4).

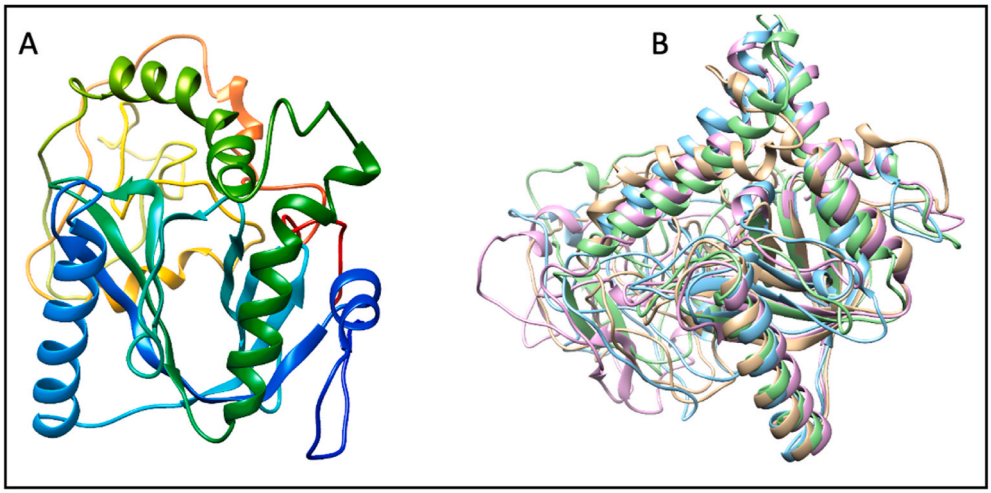

Fig. 2. 3-Dimensional structure of modeled GULP1 (A). Superimposed structures of the wild and mutant proteins (B).

Table 1
Effect of Amino acid change on *GULP1* gene and Disease association.

rsID	Mutation	PolyPhen	PhD-SNP/Score	PANTHER/Score	SNPS&GO/Score
rs1446644508	Y27C	PD	Disease/0.812	Disease/0.849	Disease/0.629
rs1357922096	G142D	PD	Disease/0.857	Disease/0.878	Disease/0.768
rs1264999716	A144T	PD	Disease/0.815	Disease/0.671	Disease/0.612
rs128246649	Y149C	PD	Disease/0.872	Disease/0.939	Disease/0.762

PD: Probably Damaging.

Table 2
Impact of Amino Acid change on the stability of *GULP1*.

rsID	Mutation	I-MUTANT	MuPro	Conservation Score	PTM Sites
rs1446644508	Y27C	Increase	Decrease	8	***
rs1357922096	G142D	Decrease	Decrease	8	Proteolytic Cleavage
rs1264999716	A144T	Decrease	Decrease	0	***
rs1282464649	Y149C	Decrease	Decrease	5	***

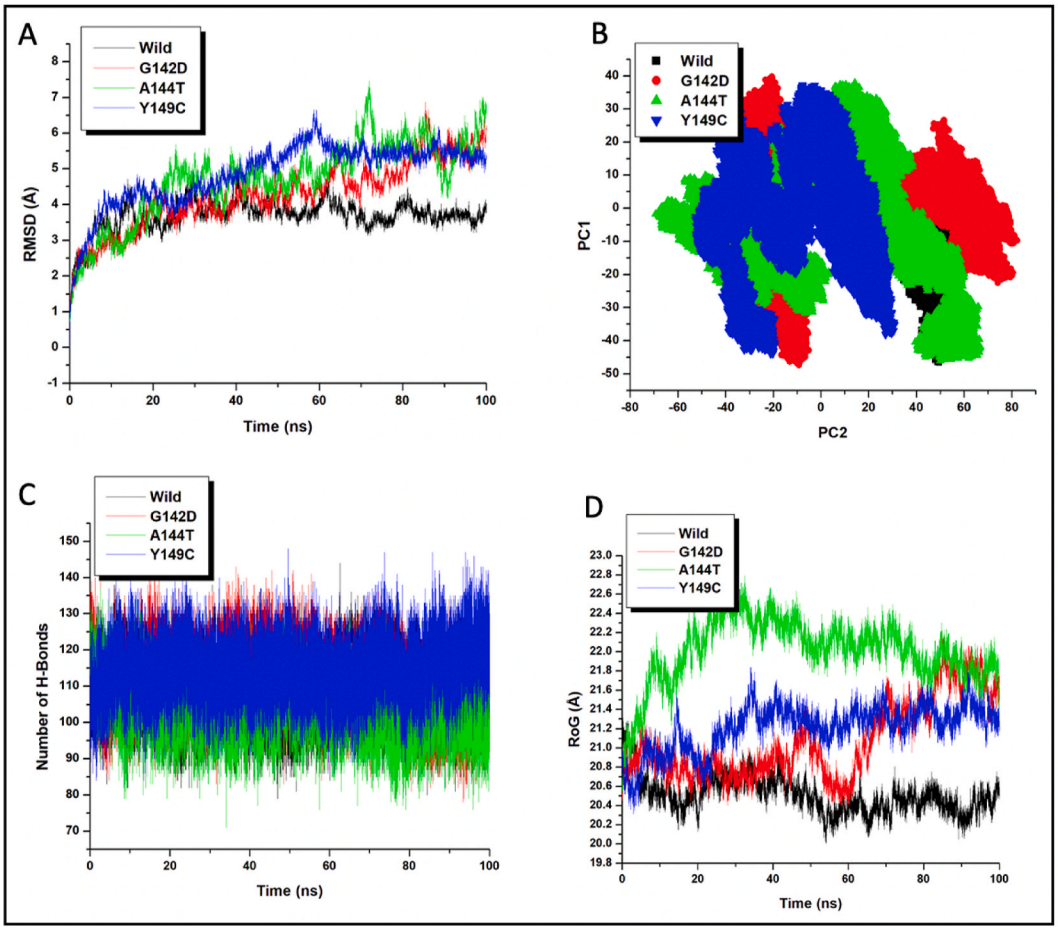

Fig. 3. Backbone RMSDs are depicted as a function of time for the wild and mutant proteins (**A**). PCA scatter plots depicting a distinct separation of motions between wild and mutant proteins (**B**). Total number of hydrogen bonds in the wild and mutant proteins (**C**). Radius of gyration of C-α atoms of the wild and mutant proteins (**D**).

The RoG estimation which measures the compactness of the Cα backbone of a protein during a production run also followed a similar trend as the RMSD plot. The wild protein exhibited high atomic compactness while the mutant proteins showed lower compactness when compared to the wild protein (Fig. 3D). The average number of hydrogen bonds in the protein before mutation was 108 while that of G142D, A144T, and Y149C were estimated to be 111, 114, and 104 respectively (Fig. 3C). The extra hydrogen bonds seen in G142D and A144T could be due to the intramolecular bonds formed by Aspartic acid and Threonine, while that seen in the Y149C protein could be as well due to bond loss when Tyrosine was mutated to Cysteine. Protein-protein interaction prediction using STRING revealed that *GULP1* interacts with 10 proteins i.e. Cell division cycle 5-like protein (CDC5L), ADP-ribosylation factor 6 (ARF6), Arf-GAP with coiled-coil (ACAP1), Multiple epidermal growth factor-like domains protein 10 (MEGF10), Engulfment and cell motility protein 3 (ELMO3), Engulfment and cell

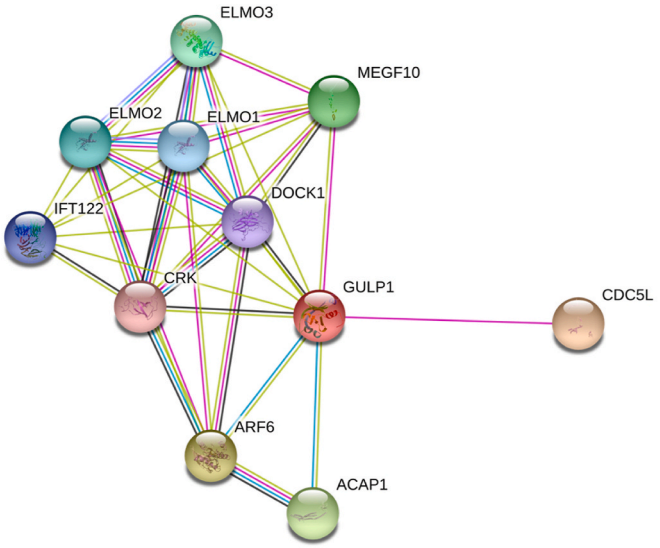

Fig. 4. Protein-Protein interaction plot of *GULP1*.

motility protein 2 (ELMO2), Engulfment and cell motility protein 1 (ELMO1), Intraflagellar transport protein 122 homolog (IFT122), Dedicator of cytokinesis protein 1 (DOCK1), and Adapter molecule crk (CRK).

4. Discussion and conclusion

Millions of SNPs have been identified within the genome, while some occur in the intronic region, others are found in the exonic region. As Genome-Wide Association studies (GWAS) increase, diverse SNPs will be identified and deposited in various databases such as dbSNP, Ensemble, gnomAD and GWAS catalogue. Due to this avalanche of SNPs data, it is somewhat becoming demanding to pinpoint particular SNPs that contribute to the onset of diseases. Computational approaches provide a scientific succour in the identification of SNPs that are pathogenic or disease-causing, these SNPs are referred to as nsSNPs. In addition, nsSNP alter protein-protein interaction, protein-DNA interaction, protein-ligand interaction, and drug metabolism. Thus, identification of nsSNPs could be used as genetic biomarkers for the diagnosis of diseases [39,40].

Each of the bioinformatics tools used in this study was developed using a different algorithm, which may lead to the possibility of having divergent prediction for the same analysis. However, the reproducibility afforded by using multiple bioinformatics tools for each analysis, ensures precision and accuracy of the results. Tools employed in this study include SIFT, PolyPhen2, PhD-SNP, Musite, I-MUTANT, ConSurf, etc. Out of the 207 identified nsSNPs, four (rs1446644508, rs1357922096, rs1264999716, and rs128246649) were jointly predicted by PhD-SNP, PANTHER, PolyPhen, and SNPs&GO to be pathogenic. However, only three nsSNPs were predicted by I-MUTANT and MuPro to have had their structural architecture affected. rs128246649 possesses a proteolytic cleavage site, while rs1446644508 and rs1357922096 are found in a highly conserved region of *GULP1*. The molecular simulation analysis as revealed by the RMSD, RoG, PCA, and number of hydrogen bonds corroborated these findings.

One of the major limitations of studies such as this is that the identified nsSNPs require experimental validation. As this study only predicts potential nsSNPs which could be pathogenic. In the course of prediction, the limitation inherent in the different tools used might influence the result and perhaps alter the accuracy of the result.

Funding

ON and SF are funded in part by the National Institutes of Health Common Fund to the H3ABioNet Project grant number 5U24HG006941-09. TC is an international training fellow supported by the Wellcome Trust grant (214205/Z/18/Z).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

NA.

References

[1] Dokholyan NV. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms in. Sci Rep 2017:1–18.
[2] Wang X, Tomso DJ, Liu X, et al. Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes. Toxicol Appl Pharmacol 2005;207:84–90.
[3] Yue P, Moult J. Identification and analysis of deleterious human SNPs. J Mol Biol 2006;356:1263–74.
[4] Giacomini KM, Brett CM, Altman RB, et al. The Pharmacogenetics Research Network: from SNP discovery to clinical drug response. Clin Pharmacol Ther 2007; 81:328–45.
[5] Su HP, Brugnera E, Van Criekinge W, et al. Identification and characterization of a dimerization domain in CED-6, an adapter protein involved in engulfment of apoptotic cells. J Biol Chem 2000;275:9542–9.
[6] Liu QA, Hengartner MO. Human CED-6 encodes a functional homologue of the Caenorhabditis elegans engulfment protein CED-6. Curr Biol 1999;9:1347–50.
[7] Smits E, Criekinge W Van, Plaetinck G, et al. The human homologue of Caenorhabditis elegans CED-6 specifically promotes phagocytosis of apoptotic cells elke Smits *, Wim van Criekinge *, Geert Plaetinck and Thierry Bogaert. Curr Biol 1999;9:1351–4.
[8] Chen X, Sun C, Chen Q, et al. Apoptotic engulfment pathway and schizophrenia. PloS One 2009;4.
[9] Qingchun H, Runyue H, LiGang J, et al. Comparison of the expression profile of apoptosis-associated genes in rheumatoid arthritis and osteoarthritis. Rheumatol Int 2008;28:697–701.
[10] Ma CIJ, Martin C, Ma Z, et al. Engulfment protein GULP is regulator of transforming growth factor-β response in ovarian cells. J Biol Chem 2012;287: 20636–51.
[11] Gurdasani D, Carstensen T, Fatumo S, et al. Uganda genome resource enables insights into Population history and genomic discovery in Africa. Cell 2019;179: 984–1002. e36.
[12] Nathan DM, Kuenen J, Borg R, et al. Translating the A1C assay into estimated average glucose values. Diabetes Care 2008;31:1473–8.
[13] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinf 2008;8: 1–8.
[14] Yang J, Roy A, Zhang Y. BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. Nucleic Acids Res 2013;41:1096–103.
[15] Yang J, Roy A, Zhang Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 2013;29:2588–95.
[16] Brylinski M, Skolnick J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. Proc Natl Acad Sci USA 2008;105: 129–34.
[17] Capra JA, Laskowski RA, Thornton JM, et al. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. PLoS Comput Biol 2009;5.
[18] Yang Z, Lasker K, Schneidman-Duhovny D, et al. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J Struct Biol 2012;179:269–78.
[19] Bateman A, Martin MJ, O'Donovan C, et al. UniProt: the universal protein knowledgebase. Nucleic Acids Res 2017;45:D158–69.
[20] Sherry ST, Ward M, Kholodov M, et al. dbSNP: the NCBI database of genetic variation29; 2001. p. 308–11.
[21] Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434–43.
[22] Hunt SE, McLaren W, Gil L, et al. Ensembl variation resources. Database 2018; 2018:1–12.
[23] Vaser R, Adusumalli S, Leng SN, et al. Protocol UPDATE SIFT missense predictions for genomes. Nat Protoc 2015;11:1–9.
[24] Adzhubei IA, Schmidt S, Peshkin L, et al. HHS Public Access 2010;7:248–9.
[25] Capriotti E, Fariselli P, PhD-SNPg. A webserver and lightweight tool for scoring single nucleotide variants. Nucleic Acids Res 2017;45:W247–52.

[26] Calabrese R, Capriotti E, Fariselli P, et al. Functional annotations improve the predictive score of human disease-related mutations in proteins. 2009.
[27] López-Ferrando V, Gazzo A, De La Cruz X, et al. PMut: a web-based tool for the annotation of pathological variants on proteins, 2017 update. Nucleic Acids Res 2017;45:W222–8.
[28] Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants from VarI-SIG 2014: identification and annotation of genetic variants in the context of structure, function and disease. BMC Genom 2016;16:1–12.
[29] Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 2005;33:306–10.
[30] Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins Struct Funct Genet 2006;62:1125–32.
[31] Glaser F, Pupko T, Paz I, et al. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics 2003;19(1):163–4.
[32] Soremekun OS, Olotu FA, Agoni C, et al. Drug promiscuity: exploring the polypharmacology potential of 1, 3, 6-trisubstituted 1, 4-diazepane-7-ones as an inhibitor of the 'god father' of immune checkpoint [Internet] Comput Biol Chem 2019;80:433–40. https://doi.org/10.1016/j.compbiolchem.2019.05.009. Available from:.
[33] Kumi RO, Soremekun OS, Issahaku AR, et al. Exploring the ring potential of 2,4-diaminopyrimidine derivatives towards the identification of novel caspase-1 inhibitors in Alzheimer's disease therapy. J Mol Model 2020;26:1–17.
[34] Berendsen HJC, Postma JPM, Gunsteren WF Van, et al. Molecular dynamics with coupling to an external bath Molecular dynamics with coupling to an external bath. J Chem Phys 2012;3684:926–35.
[35] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–41.
[36] Roe DR, Cheatham III TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. J Chem Theor Comput 2013;9:3084–95.
[37] Chau DDL, Yung KWY, Chan WWL, et al. Attenuation of amyloid-β generation by atypical protein kinase C-mediated phosphorylation of engulfment adaptor PTB domain containing 1 threonine 35. Faseb J 2019;33:12019–35.
[38] Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality, vol. 710; 2004. p. 702–10.
[39] Jayne A, Diane M, Chris C, et al. Targeted genome-wide investigation identifies novel SNPs associated with diabetic nephropathy. 2010. p. 77–82.
[40] Soremekun OS, Soliman MES. From genomic variation to protein aberration: mutational analysis of single nucleotide polymorphism present in ULBP6 gene and implication in immune response. Comput Biol Med 2019;111:103354.